NASA TT F-9433

NASA TT F-9433

BREAKDOWN IN NARROW SILICON p-n JUNCTIONS

V. K. Aladinskiy

Translation of "Proboy v uzkikh kremniyevykh p-n-perekhodakh."
Radiotekhnika i Elektronika, Vol. 10,
No. 1, pp. 102-111, 1965.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON JUNE 1965

NASA TT F-9433

BREAKDOWN IN NARROW SILICON p-n JUNCTIONS /102*

ABSTRACT

27717

Breakdown is investigated in step-type silicon p-n junctions having a width from 10^{-6} to 10^{-4} cm. It is demonstrated experimentally that in such junctions there exist simultaneously two breakdown mechanisms: tunnel and impact ionization, the interaction of which leads to sign reversal on the part of the temperature coefficient of the breakdown voltage and causes the differential impedance of the p-n junctions to vary sharply with temperature. It is established that the curve for the dependence of the differential resistance and temperature coefficient of the breakdown voltage on the breakdown voltage has a critical point, characterizing the onset of impact ionization by tunneling carriers. The threshold energy value for formation of electron-hole pairs by electrons is determined, and turns out to be 2.6 ± 0.3 ev. The experimental results are in qualitative agreement with the simple theoretical model.

Author

INTRODUCTION

Electrical breakdown in semiconductors has been the topic of a great many experimental and theoretical studies, from which it follows that the basic reasons for the abrupt increase in current carrier concentration in strong electric fields are quite clearly impact ionization and the tunnel effect.

---

*Numbers in the margin indicate pagination in the original foreign text.

By and large, the majority of experimental data relating to these processes have been obtained by the investigation of breakdown in p-n junctions.

The study of breakdown in silicon p-n junctions has shown that in such junctions with a width of 10^{-5} cm or more, breakdown is of the avalanche type (ref. 1), whereas tunnel breakdown occurs in very narrow junctions with a width of $\sim 10^{-6}$ cm (ref. 2).

The breakdown in p-n junctions with a width of 10^{-6} to 10^{-5} cm is probably caused by the combined action of the mechanisms indicated above (ref. 3), involving impact ionization of the carriers initially generated as a result of the tunnel effect in the volume of the p-n junction.

As shown by the authors of reference 3, this effect leads to a more rapid rise in the current with applied voltage. They have also shown (ref. 4) that in narrow p-n junctions carrier multiplication cannot occur unless the electron attains an energy of 2.3 eV without vacating the space charge region. This threshold energy is found to conflict with the value of 1.1 eV used in reference 5, in which a phenomenological theory is formulated in regard to a number of effects associated with impact ionization in silicon.

In reference 4, the threshold energy was determined on the basis of data relating to the multiplication of photoinjected carriers.

In the present study, we have investigated alloyed p-n junctions, which /103 are narrower than the diffused p-n junctions utilized by the authors of reference 4, where the volt-ampere curves of p-n junctions without illumination were studied.

The breakdown mechanism has been investigated and the threshold energy for formation of electron-hole pairs determined on the basis of data on the temperature dependence and differential resistance.

1. EXPERIMENTAL PART

In our work, the p-n junctions were prepared by fusing an aluminum bar 0.2 mm in diameter into silicon alloyed with arsenic having a resistivity from 0.008 to 0.15 ohm-cm. The fusion was performed in vacuum at a temperature of 720°C. Prior to fusion, the silicon ingots were oriented in the direction [111], cut into wafers, the resistivity of which was measured by the four-probe method. The wafers were then cut into crystals (dimensions 1.5 x 1.5 mm), which were subjected to a series of standard surface cleaning operations prior to fusion. Ohmic contact was produced by fusing in foil linings of alloy Au + 0.1% Sb.

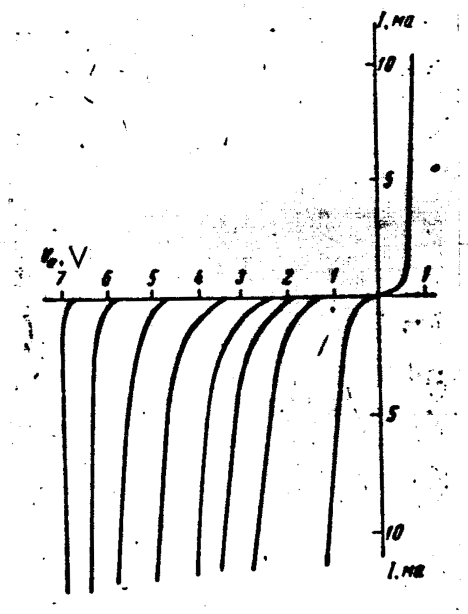

Figure 1. Typical Volt-Ampere Curves for Narrow Silicon p-n Junctions.

Typical volt-ampere characteristics of the prepared samples are shown in figure 1. The voltage on the p-n junction corresponding to a current of 10^{-2} A was arbitrarily adopted as the breakdown voltage V_B. Figure 2 shows the dependence of the breakdown voltage on the donor concentration in the base material. As apparent from the graph, a reduction in breakdown voltage with increasing

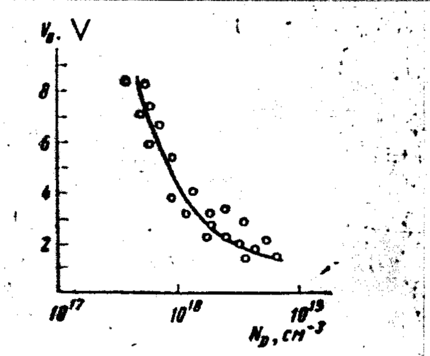

Figure 2. Dependence of Breakdown Voltage on Concentration of Alloying Impurity in Base Material.

concentration is observed up to a concentration of $\sim 10^{19}$ cm^{-3}, although at a voltage of less than 3 V a considerable spread is noted in the values of the breakdown voltage. This spread is determined by the influence of technological factors.

The capacitance was measured by the compensation method at a frequency of 300 kc with an a.c. voltage amplitude of 50 mV.

For all of the samples investigated, the graph of C^{-2} versus V_a, where C is the capacitance, V_a the external bias on the p-n junction, was a straight line whose intersection with the horizontal axis determined the contact potential difference V_i (see fig. 3). This attests to the fact that the field varies linearly with the coordinate and the concentration falls into a step distribution. The width constant W_1, i.e., the width of the p-n junction for V = 1 V, was determined by measuring the capacitance per unit area. To calculate the area, the aluminum was etched and the shape of the fused zone determined under the microscope. The calculated and measured values of the width constant were found to agree well within the limits of experimental error. /104

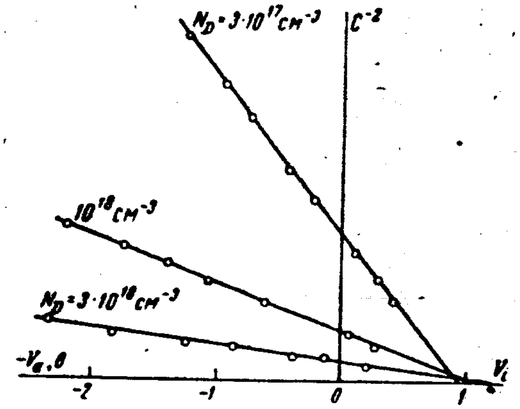

Figure 3. Dependence of Capacitance on Voltage for Various p-n Junctions.

The differential resistance $R_d = dV_B/dI|_{I=I_B}$ of the junctions was determined from the a.c. voltage appearing on the junction when supplied with a certain d.c. displacement current and 50 cps a.c. current, the amplitude of which was 1/10 the displacement current. The results are shown in figure 4.

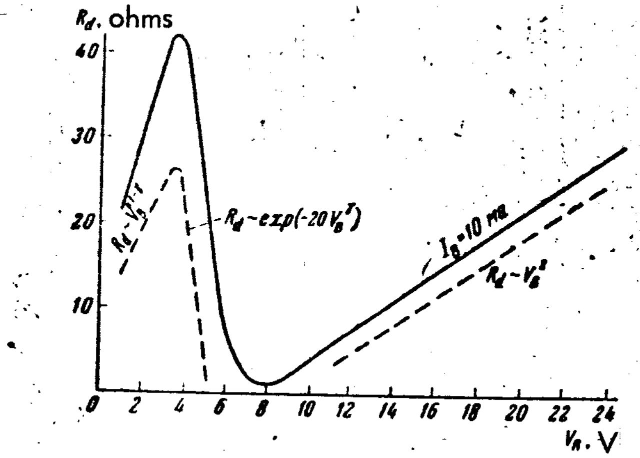

Figure 4. Dependence of the Differential Resistance on the Breakdown Voltage at the Nominal Breakdown Current of 10^{-2} A (Dashed Curve: Theoretical; Solid Curve: Experimental).

The values of the voltage temperature coefficient $\beta = dV_B/dt\,|_{I=I_B}$ and relative temperature coefficient of the differential resistance $Q = (dR_d/dT)(1/R_d)$ were determined from measurements of the breakdown voltage and differential resistance at various temperatures with the displacement current held rigorously constant by means of a direct current stabilizer, with an error of ±0.01%. The breakdown voltage was measured in a potentiometer circuit with an error of ±0.03%. The measured temperature range was -60 to +150°C; the samples /105 were placed in a thermostat, in which the temperature was regulated to within ±0.1°C.

The results are shown in figures 5 and 6.

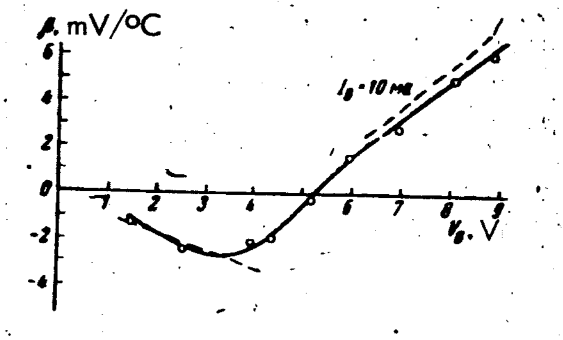

Figure 5. Dependence of the Breakdown Voltage Temperature Coefficient on the Breakdown Voltage (Dashed Curve: Theoretical; Solid Curve: Experimental).

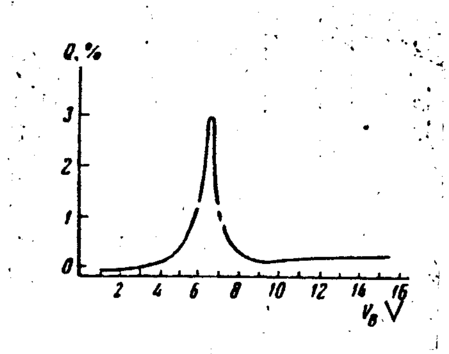

Figure 6. Dependence of the Relative Temperature Coefficient of the Differential Resistance on the Breakdown Voltage.

2. DISCUSSION OF THE RESULTS

Tunnel Breakdown in Narrow Silicon p-n Junctions

The experimental data indicate that the structure of the p-n junction satisfies the following relations:

$$\left.\begin{aligned} &F(x)=F_M\left(1-\frac{x}{W}\right),\\ &F_M=2V/W, \quad W=W_1\sqrt{V},\\ &W_1=\sqrt{\frac{\varepsilon}{2\pi q}\frac{(N_A+N_D)}{N_AN_D}}=a\sqrt{\frac{N_A+N_D}{N_AN_D}},\\ &a=\sqrt{\varepsilon/2\pi q,}\end{aligned}\right\} \qquad (1)$$

where F_M is the maximum field strength, W is the width of the junction, $V = V_\alpha + V_i$ is the total voltage on the junction in the reverse direction, V_α is the supplied voltage, V_i is the contact potential difference, W_1 is the width constant, ϵ is the dielectric constant, q is the electronic charge, N_A and N_D are the surplus acceptor and donor concentrations in the p- and n-type regions, respectively.

Assuming that breakdown is attributable to the tunneling of electrons from the valence band into the conduction band, in the one-dimensional case it is not difficult to obtain an approximate expression for the current through the p-n junction:

$$I \simeq (qSz/d^3)f(F_M)\nu(F_M)W, \qquad (2)$$

where z is the number of valence electrons in the elementary cell, d is the lattice period, f is the penetration probability, $\nu = qFd/h$ is the oscillation frequency within a single band, h is the Planck constant, S is the area of the junction.

In equation (2), we neglect the nonuniformity of the field in the p-n junction, supposing the field to be everywhere equal to its maximum. This means that carrier tunneling occurs for the most part in the region of maximum field /106 strength, the size of which is considerably less than the total width of the p-n junction, so that the field nonuniformity in this region may be neglected.

The field distribution in the junction will be taken into account below as a function of the maximum field F_M due to the applied voltage V, as given by equations (1).

As shown in reference 2, tunneling occurs in silicon with the participation of phonons. In this connection, the penetration probability f can be written (ref. 6) as

$$f \simeq \exp\left[-\frac{4\sqrt{2m^*}}{3q\hbar F}(E_g-\hbar\omega)^{3/2}\right] = \exp\left[-\frac{A(T)}{F}\right], \qquad (3)$$

where m^* is the effective mass, E_g is the width of the forbidden band, $\hbar\omega$ is the phonon energy, $\hbar$ is the Planck constant divided by 2π.

The quantity A(T) is the expression contained in the exponent and depends implicitly on the temperature. Consequently, equation (2) can be written in the form

$$I \simeq (zq^2 S/d^3) F_M W \exp\left[-A(T)/F_M\right]. \qquad (4)$$

Taking equations (1) into account and differentiating (4) as an implicit function, we find the differential resistance

$$R_d = V/I\left[1+\frac{A(T)}{2F_M}\right]. \qquad (5)$$

Substituting the value of the effective mass $m^* = 0.36\ m_0$, where m_0 is the free electron mass, i.e., the reduced electron-hole mass, into the expression for A(T) (3), we obtain a value of $A \simeq 4 \cdot 10^7$ V/cm. Inasmuch as $F_M \sim 10^6$ V/cm, we finally get

$$R_d \simeq 2F_M V / IA(T). \qquad (6)$$

Since it is primarily the exponential factor of (3) that varies in breakdown, for the given field distribution,

$$\ln R_d = D + B(V_a + V_i)^{-1/2}, \qquad (7)$$

where D and B are constants for a given junction.

Consequently, for step-type p-n junctions in the case of tunnel breakdown, the differential resistance must decrease as $\sim I^{-1}$, and its logarithm is a linear function of the quantity $(V_\alpha + V_i)^{-1/2}$.

The relations obtained are in harmony with the experimental data (fig. 7).

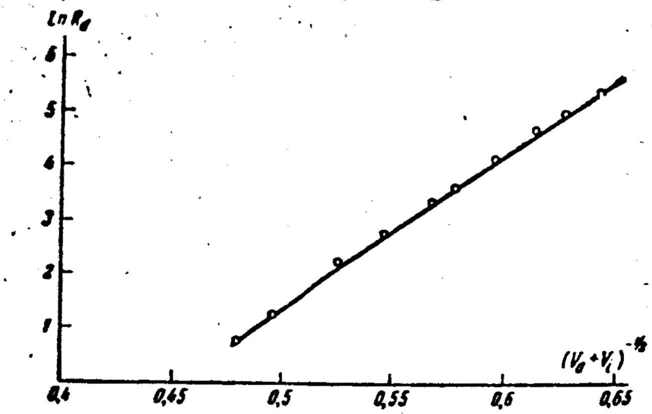

Figure 7. Typical Characteristic for p-n Junction with Tunnel Breakdown.

Of considerable interest is the characteristic of relatively sharp delineation of the breakdown in p-n junctions with different widths and, consequently, with different breakdown voltages V_B.

As the experiment has shown (see fig. 4), when V_B increases, i.e., when the width of the p-n junction increases, the quantity R_d measured at the same value of the breakdown current, increases. This increase is explained by a decrease in the maximum field strength when the p-n junction is widened.

The dependence of the differential resistance R_d on the breakdown voltage V_B (assumed purely arbitrarily) is given by equation (6), where the connection between the maximum field in the junction F_{MB} and breakdown voltage V_B is determined empirically:

$$F_{MB} = F_0/V_B^{\gamma}, \qquad (8)$$

$F_0 = 2\cdot 10^6$ V/cm, $\gamma = 0.25$.

Based on equation (8), the expression (6) transforms to /107

$$R_d = 2F_0V_B^{1-\gamma}/I_BA(T). \qquad (9)$$

The dashed curve in figure 4 was constructed on the basis of equation (9) for $I_B = 10^{-2}$ A.

The temperature characteristics of tunnel breakdown are readily obtained by assuming that in the investigated temperature range (-60 to +150°C) thermal expansion of the bodies is the principal source of the temperature dependence of the penetration probability f, i.e.,

$$\frac{dA}{dT}\frac{1}{A} \simeq -4\cdot 10^{-4}.$$

Allowance for the influence of phonons dces not play a decisive part in the present case, for all that counts is the sign of the temperature dependence. Then for the temperature coefficient of the breakdown voltage (β) we obtain (see fig. 5)

$$\beta \simeq \frac{dV_B}{dT}\bigg|_{I=I_B} = 2\frac{dA(T)}{dTA(T)}V_B \simeq -8\cdot 10^{-4}V_B. \qquad (10)$$

Similarly, the relative temperature coefficient of the differential resistance (see fig. 6) is

$$Q = dR_d/dTR_d = -2\cdot 10^{-4}. \qquad (11)$$

Consequently, the results obtained are found in qualitative agreement with the experimental data.

Impact Ionization in Wide Silicon p-n Junctions

The current in the region of large reverse displacement is given by the expression (ref. 7)

$$I = I_0(F, T)M, \qquad (12)$$

where I_0 is the current due to heat generation, M is the multiplication factor. The quantity $I_0(F, T)$ has two components: the generation-recombination current and diffusion current, the relative proportion of each depending on the applied bias and temperature. Assuming that there are traps on one type situated in /108 the middle of the forbidden band, we obtain for the current

$$I_0(F, T) = S\left(\frac{qn_iW}{2\sqrt{\tau_{p0}\tau_{n0}}} + \frac{qn_i^2L_p}{N_D\tau_{p0}}\right), \qquad (13)$$

where L_p, τ_{p0} are the diffusion length and lifetime of holes in a strongly alloyed n-type material, τ_{n0} is the electron lifetime in a strongly alloyed p-type material, n_i is the current carrier concentration in the intrinsic semiconductor.

Allowing for simplicity that the properties of the electrons and holes are identical, we have for the multiplication factor (ref. 7)

$$M = 1\Big/\left(1 - \int_0^W \alpha(F)dx\right). \qquad (14)$$

It follows from reference 8 that the impact ionization coefficient can be written as

$$\alpha(F) = \alpha_0(F)\exp(-b^2/F^2), \qquad (15)$$

where $\alpha_0(F)$ is a function depending much less on the field than the exponential factor, b is some characteristic field in which the mean carrier energy becomes of the same order of magnitude as the ionization energy $\mathcal{E}_i$. Inasmuch as the impact ionization coefficient is a sharply varying function of the field, the integral in equation (14) can be replaced by the maximum of the integrand function, i.e.,

$$M = 1 / [1 - p(W) W \alpha_0(F_M) \exp(-b^2 / F_M^2)], \tag{16}$$

where $0 < p(W) < 1$. If in the case of the tunnel effect the breakdown voltage is a fairly arbitrary concept, for the impact ionization case we formally introduce the breakdown condition ($M \longrightarrow \infty$)

$$\int_0^{W_B} \alpha(F) dx = 1. \tag{17}$$

It can be shown that this condition leads to a transcendental equation for the breakdown voltage V_B and donor concentration N_D, the approximate solution of which has the form

$$V_B = C N_D^{-\mu}, \tag{18}$$

where $\mu = 0.62$; C is a constant. Hence the maximum field at breakdown is

$$F_{MB} = F_0 V_B^{-\gamma}; \ \gamma = (1 - \mu) / 2\mu, \tag{19}$$

an expression which coincides with equation (7), which was derived empirically for tunneling.

Differentiating equation (12) as an implicit function, we readily obtain an expression for the differential resistance as a function of breakdown voltage, making use of equation (18):

$$R_d = (4 V_B^2 q n_i S / I_B^2 b^2 a^2 \tau_0)(n_i L_p + (a/2) V_B^{-\gamma} C^{1/\mu}) \simeq V_B^2 / I_B^2, \tag{20}$$

where $\tau_0 = \tau_{p0} = \tau_{n0}$.

Consequently, the quantity R_d practically increases as $\sim V_B^2$ for I_B = const. Figure 4 shows the dependence given by equation (20). Clearly, the agreement with experiment is fairly good.

The temperature dependence of the breakdown voltage is governed largely /109 by the variation in the parameter b with temperature. Assuming that all collisions result in energy loss on the part of the electron, which is a good approximation for silicon at temperatures $\sim 300°K$, an expression can be obtained for the breakdown voltage temperature coefficient β. In fact, we can write on the basis of reference 8

$$b = b_0 \operatorname{cth}^{1/2}(\hbar\omega_0 / 2kT), \tag{21}$$

where ω_0 is the optical phonon frequency.

Making use of equation (21) and the breakdown condition (17), we obtain

$$\beta = dV_B/dT = \hbar\omega_0 b_0^2 a^2 V_B^{1/\mu} \Big/ 2kT^2 \operatorname{sh}^2 \frac{\hbar\omega_0}{2kT} 2C^{1/\mu}\left[1 + \frac{b^2a^2}{2C^{1/\mu}} V_B^{(1-\mu)/\mu}\right]. \tag{22}$$

Figure 5 illustrates this dependence for T = 300°K.

It follows from equations (22) and (20) that $Q = dR_d/dTR_d$ is positive, increasing slightly with the breakdown voltage (see fig. 6).

Transition Region; Determination of the Threshold Energy for Electron-Hole Pair Production

Consider a step-type p-n junction (fig. 8) in which $N_A \gg N_D$ and the total space charge lie in the n-type region.

As already noted above, the generation of carriers proceeds mainly in the region of maximum field strength, the size of which is $\sim E_g/qF_m$ = 50 to 100 Å. The electrons falling into the conduction band are accelerated into regions of fairly strong field and can ionize, provided the required energy is obtained.

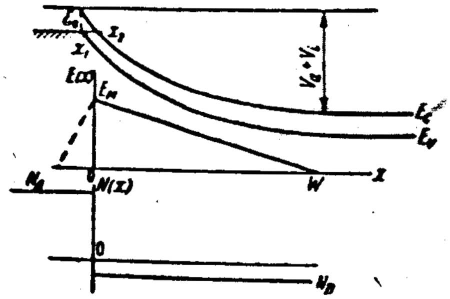

Figure 8. Model of a Step-Type p-n Junction.

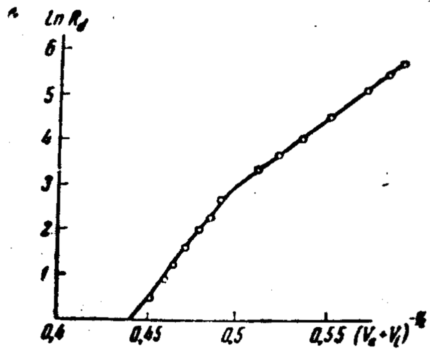

Figure 9. Characteristic of a p-n Junction with Impact Ionization by Tunneling Carriers.

Clearly, the energy derived from the field is

$$qV_F = q(V_a + V_i) - (E_g - \hbar\omega). \tag{23}$$

Let us assume that multiplication as given by equation (12) is applicable in the given case. This is valid if the distance at which carrier multiplication is induced by impace ionization is greater than the mean free path. Inasmuch as $\lambda \sim 100$ Å, we have $W/\lambda \simeq 10$.

The value of the current will be summed from the thermal current (13) and the tunnel current (2).

Carrying out computations analogous to those above, we obtain /110

$$R_d = \frac{4V_B^2 q n_i S}{I_B^2 b^2 a^2 \tau_0}\left[n_i L_p + \frac{a}{2} C^{1/2\mu} V_B^{-\gamma} + \frac{zq}{hd^2} F_M W_B \exp\left(-\frac{AV_B^\gamma}{F_0}\right)\right] \simeq$$
$$\simeq \frac{1}{I_B^2} \exp\left(-\frac{AV_B^\gamma}{F}\right). \tag{24}$$

Analyzing equation (24), it may be asserted that the behavior of the differential resistance in the transition region is dictated largely by the behavior of the function $I_0(F, T)$.

Inasmuch as the tunnel current exceeds the thermal current in the transition region, the differential resistance measured for different junctions at some constant value of the breakdown current I_B will decrease with increasing breakdown voltage V_B, since with diminishing field strength the tunnel current falls off as $\sim \exp(-A/F_{MB}) = \exp(-AV_B^{\gamma}/F_0)$. In other words, the condition $I_B = I_0$, $M = \text{const}$ means that measurements of R_d in the transition region correspond to steeply rising values of the multiplication factor.

The most abrupt characteristic in the transition region is the temperature dependence of the differential resistance (see fig. 6).

As the temperature is increased, the carrier penetration probability rises sharply, and breakdown will be governed largely by the tunnel effect, but since the field strength decreases with increasing width of the p-n junction according to (9), R_d will increase.

The breakdown voltage temperature coefficient β, of course, has a zero point in the transition region (see fig. 5). Here the maximum negative β, like the maximum R_d, signals the inception of impact ionization by tunneling carriers, and for the same given p-n junction they coincide.

It proved very instructive, in this connection, to investigate p-n junctions for which the two breakdown mechanisms could be sharply distinguished.

In fact, for such p-n junctions the graph of ln R_d versus $(V_\alpha + V_i)^{-\frac{1}{2}}$ (see /111 fig. 9) disclosed an abrupt bend at the point where the onset of impact ionization was observed. This made it possible to determine, directly from the position of the bend, the energy acquired by the carriers from the field, since the contact potential V_i was determined from measurements of the capacitance.

For the energy derived from the field we have

$$qV_F = qV_a \simeq \mathcal{E}_i + (\hbar\omega_0/\lambda)W, \tag{25}$$

since $V_i = E_g - \hbar\omega$. In this way, it is possible to determine the total potential at which multiplication begins in a given p-n junction. In order to ascertain the threshold energy for electron-hole pair production we must eliminate from (25) the losses due to collisions with phonons. To do this, we constructed the dependence of the potential across the junction for which $M \simeq 1$ on the junction width W_0 for p-n junctions with various widths. i.e., for those p-n junctions disclosing a bend.

Approximating this dependence to zero, we directly obtain the threshold energy for electron-hole pair production by electrons, since electrons tunneled directly into the p-n junction.

The resultant value of 2.6 ± 0.3 eV agrees with the value of 2.3 ± 0.1 eV obtained in reference 4.

The slope of the curve in figure 10 characterizes a certain phonon slowing-down field $\hbar\omega_0/\lambda$ caused by the generation of optical phonons. For an optical phonon energy $\hbar\omega_0 = 0.063$ eV, the value determined for the mean free path in breakdown is λ = 60-70 Å.

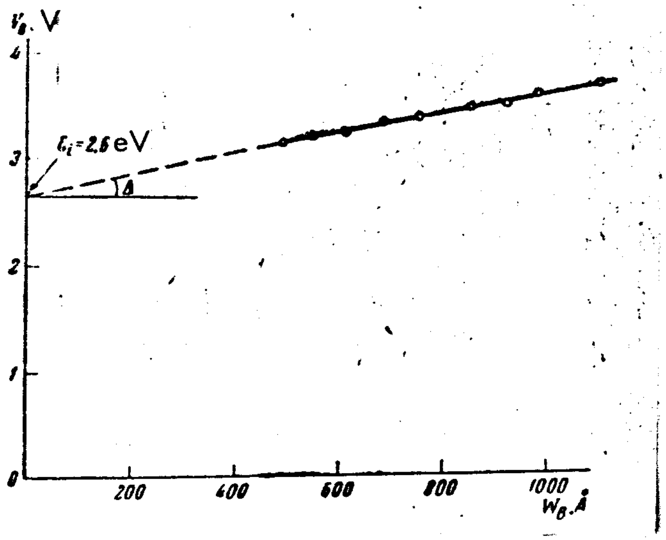

Figure 10. Dependence of the Voltage Across the p-n Junction for Which $M \simeq 1$ on the Width of the Junction ($\tan \Delta = h\omega_0/q\lambda \simeq 10^5$ V/cm; $\hbar\omega_0 = 0.063$ eV; $\lambda = 60$ to 70 Å).

It must be noted, in conclusion, that the real structure of the p-n junction was not borne in mind in the present work, including statistical and stochastic fluctuations of the impurity, the presence of structural defects.

The influence of these factors is especially noticeable in the prebreakdown region and will be less marked in deep breakdown with the high current densities at which our experiments and calculations were carried out. Taking these factors into account leads to a quantitative correction, but the results obtained in the qualitative analysis of the effects should remain in force.

The author takes this opportunity to express his gratitude to B. M. Vul and L. V. Keldysh for a number of invaluable comments.

REFERENCES

1. McKay, K. G. Phys. Rev., Vol. 94, p. 977, 1954.

2. Chynoweth, A. G., Feldman, W., Zee, C, Zogan, R. and G. Pearson. Phys. Rev., Vol. 118, p. 425, 1960.

3. Chynoweth, A. G. and K. G. McKay. Phys. Rev., Vol. 106, p. 418, 1957.

4. Chynoweth, A. G. and K. G. McKay. Phys. Rev., Vol. 108, p. 29, 1957.

5. Shockley, W. Problems Related to p-n Junctions in Silicon (paper presented at the International Conference on Semiconductor Physics, Prague, Czechoslovakia, 1960). Czech. J. Phys., Vol. B11, p. 81, 1961.

6. Keldysh, L. V. Zh. Eksperim. i Teor. Fiz., Vol. 34, p. 962, 1958. English translation: Soviet Physics - JETP, Vol. 34, p. 665, 1958.

7. Chin-Tang Sah, Noyce, R. and W. Shockley. Proc. IRE, Vol. 45, p. 1228, 1957.

8. Keldysh, L. V. Zh. Eksperim. i Teor. Fiz., Vol. 37, p. 713, 1959. English translation: Soviet Physics - JETP, Vol. 37, p. 509, 1960.

Received by the editors of Fizika Tverdogo Tela,
9 August 1963

Received by the editors of Radiotekhnika i Elektronika,
4 October 1963

Translated for NASA by Stemar Engineering Inc.
4940 Long Beach Blvd., Long Beach, California.